UBIQUITOUS KNOWLEDGE PROCESSING

Bachelor/Master Thesis

TECHNISCHE UNIVERSITÄT DARMSTADT

Question Answering over Scientific Contents

Motivation

A very important step in conducting research is reviewing existing literature. This allows to develop novel ideas, identify current gaps and eventually produce impactful research. However, the number of publications grows rapidly and is far too big for humans to study in detail. In order to overcome this challenge, innovative tools that allow to automate part of the literature review process have to be developed. One such technology is QA, where the answer to a question is automatically produced based on scientific background knowledge. This task is challenging due to scarce data, complex nature of the questions and underlying texts, and long documents.

Task Description

- Design, implement and evaluate novel models for Question Answering that deal with the special challenges of QA over scientific literature (i.e. specialized language, few-shot learning, long documents)
- Tasks of the entire QA pipeline can be worked on, e.g. information retrieval, span extraction, answer generation, etc.

References

- Jin, Qiao, et al. "Biomedical Question Answering: A Comprehensive Review." arXiv preprint arXiv:2102.05281 (2021).
- Dasigi, Pradeep, et al. "A Dataset of Information-Seeking Questions and Answers Anchored in Research Papers." arXiv preprint arXiv:2105.03011 (2021).

Analysis	■■■■□
Programming	■■■■■
Literature	■■■□□

Contact

Prof. Dr. Iryna Gurevych

Tim Baumgärtner

thesis@ukp.informatik.tu-darmstadt.de

www.ukp.tu-darmstadt.de